Geophysical Journal International

Geophys. J. Int. (2018) **215,** 942–958
Advance Access publication 2018 July 28
GJI Seismology

doi: 10.1093/gji/ggy316

Using surface waves recorded by a large mesh of three-element arrays to detect and locate disparate seismic sources

Wenyuan Fan,[1] Catherine D. de Groot-Hedlin,[2] Michael A. H. Hedlin[2] and Zhitu Ma[3]

[1]*Department of Geology and Geophysics, Woods Hole Oceanographic Institution,* 266 *Woods Hole Rd., Woods Hole, MA* 02543, *USA. E-mail: wfan@whoi.edu*
[2]*Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA, 92093, USA*
[3]*Department of Earth, Environmental and Planetary Sciences, Brown University, 324 Brook St., Providence, RI, 02912, USA*

Accepted 2018 July 27. Received 2018 July 13; in original form 2018 March 26

SUMMARY

Surface waves recorded by global arrays have proven useful for locating tectonic earthquakes and in detecting slip events depleted in high frequency, such as glacial quakes. We develop a novel method using an aggregation of small- to continental-scale arrays to detect and locate seismic sources with Rayleigh waves at 20–50 s period. The proposed method is a hybrid approach including first dividing a large aperture aggregate array into Delaunay triangular subarrays for beamforming, and then using the resolved surface wave propagation directions and arrival times from the subarrays as data to formulate an inverse problem to locate the seismic sources and their origin times. The approach harnesses surface wave coherence and maximizes resolution of detections by combining measurements from stations spanning the whole U.S. continent. We tested the method with earthquakes, glacial quakes and landslides. The results show that the method can effectively resolve earthquakes as small as ∼M3 and exotic slip events in Greenland. We find that the resolution of the locations is non-uniform with respect to azimuth, and decays with increasing distance between the source and the array when no calibration events are available. The approach has a few advantages: the method is insensitive to seismic event type, it does not require a velocity model to locate seismic sources, and it is computationally efficient. The method can be adapted to real-time applications and can help in identifying new classes of seismic sources.

1 INTRODUCTION

Detecting seismic sources is one of the most fundamental goals in seismology. Catalogues of seismic events not only provide important constraints on fault zone structures, earthquake physics, and earthquake triggering and interactions (e.g. Douglas 1967; Jordan & Sverdrup 1981; Thurber 1983, 1985; Waldhauser & Ellsworth 2000; Hauksson & Shearer 2005; Shearer *et al.* 2005; Lin *et al.* 2007; Waldhauser & Schaff 2008; Shelly & Hardebeck 2010; Lin 2013; Matoza *et al.* 2013; Fan & Shearer 2016; Cattania *et al.* 2017), but also serve as the base to image Earth structure and to investigate tectonics (e.g. Zhu *et al.* 2012; Ma & Masters 2015; Gibbons *et al.* 2017). Therefore, complete catalogues with low uncertainties are highly desirable. Traditional catalogues are built upon detecting and locating seismic events with short-period body wave arrival times (*P* and/or *S* waves). Such analyses involve determinations of seismic phases and constructions of inverse problems using an assumed velocity model (e.g. Thurber 1983, 1985; Waldhauser & Ellsworth 2000; Lin & Shearer 2006; Trugman & Shearer 2017). Recently, matched filter and other advanced machine learning techniques have been widely applied for seismic event detections due to advances in computing resources, which do not require accurate knowledge of the subsurface velocity structure (e.g. McGuire 2008; Meng & Peng 2014; Huang & Beroza 2015; Yoon *et al.* 2015; Li *et al.* 2018). However, a relatively complete event template library is necessary to assure the performance of the matched filter algorithms (e.g. McGuire 2008; Meng & Peng 2014; Huang & Beroza 2015).

Array analysis provides another means for seismic source detections. Both small aperture and global arrays have been successfully implemented targeting a wide range of seismological and geophysical problems (e.g. Ringdal & Husebye 1982; Ruud *et al.* 1988; Hedlin *et al.* 1991; Shearer 1994; Gibbons *et al.* 2005; Gibbons & Ringdal 2006; Ekström 2006; Gibbons *et al.* 2010; Chen *et al.* 2011). Perhaps the most exciting findings from these detection algorithms are identifications of new classes of seismic events that were missed in most of the global earthquake catalogues (Ekström *et al.* 2003, 2006; Tsai & Ekström 2007; Nettles & Ekström 2010; Chen *et al.* 2011). For instance, Ekström *et al.* (2003) used intermediate-period surface waves to detect and locate seismic events associated with glacier displacement processes in multiple regions. The magnitudes of these glacial quakes are equivalent to those of ∼M5 earthquakes, but they were absent from traditional earthquake catalogues because they lacked high-frequency seismic radiations (Tsai & Ekström 2007; Chen *et al.* 2011). These events are directly associated with glacial calving and ice flow (Joughin *et al.* 2008; Nettles *et al.* 2008; Tsai *et al.* 2008; de Juan *et al.* 2010; Nettles & Ekström 2010; Chen *et al.* 2011; Murray *et al.* 2015a,b). Therefore, it is important to monitor the occurrence of these seismic events because

of their broad implications to global warming and sea level rise (e.g. Veitch & Nettles 2012; Olsen & Nettles 2017). Often these surface wave detectors are based on a known velocity model, which may introduce uncertainties in event locations due to possible inaccuracies of the model. In addition, deconvolution and the Hilbert transform (envelope function) are applied to seismic data to extract the source signals and to neutralize the radiation pattern (Ekström 2006). These procedures are necessary for using globally distributed stations to constructively interfere the signals. However, deconvolution can be computationally intensive, and the Hilbert transform does not use the polarity information from the data.

Here, we experiment with detecting seismic sources with surface waves at 20–50 s period that were recorded by large aperture arrays. By dividing a large aperture array into Delaunay triangles, our approach takes advantage of the highly coherent surface waves between adjacent stations. The coherent waveforms can be directly used for detections, and the Hilbert transform is not necessary for our approach. This allows us to achieve a greater spatial and temporal resolution of the source location. In addition, our approach does not require a velocity model nor knowledge of the seismic event type. The method can be routinely and automatically applied to multiple large arrays simultaneously to enhance the spatial resolution, has the potential to be implemented in real time, and can potentially help detect and locate new classes of earthquakes. This paper focuses on introducing the algorithm, empirically exploring its associated uncertainties, and highlighting its potential to detect non-earthquake sources, for example glacial quakes and landslides. Results of comprehensive implementations of the method to the U.S. and globally distributed arrays will be presented in a follow-up paper.

2 DATA AND METHOD

2.1 Data: U.S. continental arrays

We use stations registered at the International Federation of Digital Seismograph Networks within 6000 km of 39.8°/ − 98.6°, which is taken as a proxy for the centre of the USArray Transportable Array (Busby *et al.* 2006). We then download continuous vertical-component long-period channel (LHZ) data of the accessible stations from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS). The LHZ data are digitized at one sample per second. The stations are primarily from the USArray Transportable Array, which has a nominal station spacing of ∼70 km (Busby *et al.* 2006). In addition, several PASSCAL flexible arrays, the Pacific Northwest Seismic Network (PNSN), the Southern California Seismic Network (SCSN) and a few Canadian and Mexican stations are used for analysis as well. Depending on the time, available networks may vary. For instance, there are 852 available stations on 2010 April 4 (Fig. 1).

In this study, we downloaded data of 2010 April as an example month. We truncate the records into segments of 25 hr. Each segment starts at 12 a.m. of each day (in UT) and ends at 1 a.m. the next day to avoid missing events recorded around mid-night. All the records of the time period were organized into one data file to be used for later analysis. We do not remove instrument responses, but bandpass filter the records at 20–50 s period (0.02–0.05 Hz) taking advantage of the high signal-to-noise ratio (SNR) in the frequency band (Shearer 1994; Ekström 2006; McGuire 2008). The high SNR signals, which are dominated by Rayleigh waves, allow teleseismic distance moderate size earthquakes to be clearly observed across the arrays.

2.2 Method

The proposed method is based on the AELUMA (Automated Event Location Using a Mesh of Arrays) method introduced in de Groot-Hedlin *et al.* (2014) and de Groot-Hedlin & Hedlin (2015). de Groot-Hedlin *et al.* (2014) first developed a detection method for gravity waves. de Groot-Hedlin & Hedlin (2015) and de Groot-Hedlin & Hedlin (2018) later extended the method for not only detections but also locations, which have been successfully applied to infrasound signals and earthquake body waves (de Groot-Hedlin & Hedlin 2015; de Groot-Hedlin *et al.* 2017; de Groot-Hedlin & Hedlin 2018). These methods divide larger arrays into small subarrays, group the detected signals, and then form inverse problems to locate the sources. However, due to the different natures of the targeted signals (incoherent versus coherent waves between adjacent stations), the signal processing details are quite different in practice. In this study, we first divide the large arrays into small subarrays, each comprising three stations. Each subarray is called a triad (Fig. 1). Second, tau–p analysis is applied to detect signals, and the detections are screened through quality control. Third, the detections are grouped into non-overlapping clusters. Fourth, detections of each cluster are used to locate one seismic source. Finally, the quality of each located seismic event is assessed to avoid duplicates and a catalogue is populated with the located events.

2.2.1 Stage 1: array discretization

All the arrays are discretized into non-overlapping triad subarrays (Fig. 1) using Delaunay triangulation (Lee & Schachter 1980; Thompson & Shure 2016). Each triad is designed to detect coherent planer signals travelling across the subarray, and to resolve the signal arrival time, speed and direction. For optimal performance only triads with side lengths between 10 and 600 km and interior angles between 30° and 120° are used. By discretizing large arrays into basic triad elements, we can effectively combine multiple large arrays despite the fact that they are spatially disconnected. For example, there are 1030 triads formed with the available stations in the study area on 2010 April 4 (Fig. 1). The median of the side length for this array configuration is 75.6 km (insert Fig. 1), which makes the triads well suited for detecting planer wave signals. Such discretization is performed daily to adapt to the evolving array configuration, for example USArray moved from the West Coast to the East Coast.

2.2.2 Stage 2: signal detection

We use tau–p array processing to detect coherent signals crossing each triad (e.g. Havskov & Ottemöller 2010; Hedlin *et al.* 2018). For each triad, centroid arrival time, direction of propagation and local phase velocity can be determined from time delays among three stations computed from cross-correlations of the waveforms (Fig. 2). For a triad with a wave front propagating horizontally across the subarray (Fig. 2a), the wave propagation can be parametrized with a phase velocity $\underline{v}$ (or slowness $\underline{p}$) and propagation direction θ, where θ is clockwisely measured from the north. A set of time delays, $T_{i,j}$ between station i and j ($i \neq j$ and $i, j \in [1, 2, 3]$), can be obtained by cross-correlating signals recorded at each station. Ideally, a coherent signal would cause the sum of the delay times to be zero ($t_{\mathrm{sum}} = T_{1,2} + T_{2,3} + T_{3,1} = 0$). However, surface wave dispersion,

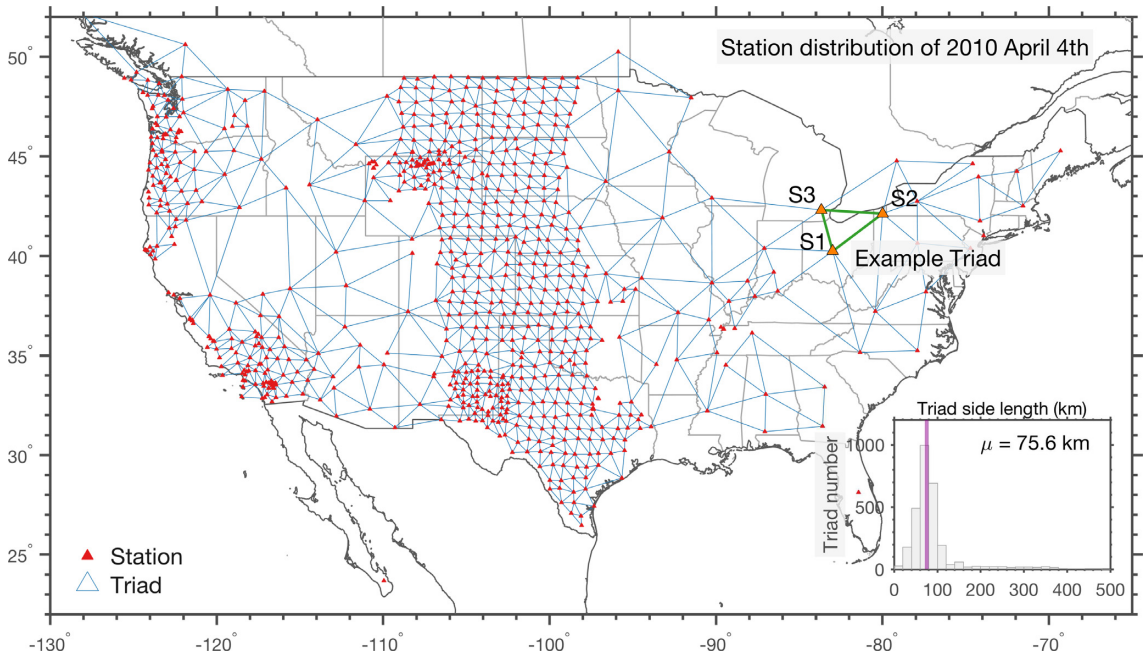

Figure 1. Station distribution and triads on 2010 April 4. The network includes USArray Transportable Array, several PASSCAL flexible arrays (FA), the Pacific Northwest Seismic Network (PNSN), the Southern California Seismic Network (SCSN), and a few Canadian and Mexican stations. Insert: triad side length distribution. The median triad side length, μ, is 75.6 km for 2010 April 4.

waveform distortion and noise will lead to some inconsistency ($t_{\text{sum}} = T_{1,2} + T_{2,3} + T_{3,1} \neq 0$). Therefore, we empirically require $|t_{\text{sum}}| \leq t_{\text{res}}$ to accept a useful number of detections before removing outliers (Cansi 1995). Here, we empirically set t_{res} as 60 s. These values ($T_{i,j}$) are used to compute a beamformed signal corresponding to a point at the triad centroid. For a coherent signal, we use the tau–p method to calculate the direction of propagation and local phase velocity (signal slowness) of the coherent signal:

$$\begin{bmatrix} T_{1,2} \\ T_{2,3} \\ T_{3,1} \end{bmatrix} = \begin{bmatrix} \underline{r}_2 - \underline{r}_1, & \underline{r}_3 - \underline{r}_2, & \underline{r}_1 - \underline{r}_3 \end{bmatrix}^T \cdot \underline{p} \tag{1}$$

where $\underline{p} = \begin{bmatrix} p_x, & p_y \end{bmatrix}^T$ and $\underline{r}_j - \underline{r}_i$ is the distance between station i and j. The least-squares solution of $\underline{p}$ is used to determine

$$\theta = \tan^{-1}\left(\frac{p_x}{p_y}\right) \tag{2}$$

and

$$|\underline{v}| = \sqrt{\frac{1}{p_x^2 + p_y^2}} \tag{3}$$

Here, centroid arrival times, direction of propagations and local phase velocities are independently measured at each triads across the whole array. This is different from Cansi (1995), which focuses on seismic phase detection, and only estimates one propagation direction and phase velocity across the whole array.

In practice, the 25 hr long waveforms are broken into 600 s segments with a 180 s time-step between windows. For each 600 s time window, we cross-correlate the three traces of the triads with each other assuming planer wave propagation (Figs 2 c–e). The maximum delay time ($T_{i,j}$) is constrained to be no larger than the direct traveltime along geodesic minimal arc, or great-circle path, between two stations, assuming a phase velocity of 2 km s^{-1}. For the example triad (Fig. 2a), the mean, median and standard deviation of t_{sum} for all the measurements are 0.82, 0.86 and 0.14 s for the three detections in 2010 April 4, suggesting the signals are highly coherent. In addition, the average of the three cross-correlation coefficients within a triad has to exceed 0.5 to accept the measurement of a given time window as a candidate for later analysis. From the tau–p analysis, the beam power is given by the maximum amplitude of the beamformed signal and the centroid arrival time is identified as the time at which this maximum occurs (Figs 2 a and b). Detections with beam power amplitude less than 5 counts are removed. The local phase velocity is required to be within 2.5–5 km s^{-1} to assure the detected signals are surface waves, otherwise, the measurements of the time window will be discarded. For each triad, we further remove the duplicate measurements with the centroid arrival times within 300 s, and only keep the measurements with the highest beam power (maximum of linearly stacked waveforms). As an example, Fig. 2(b) shows the remaining measurements of 2010 April 4 at the example triad (Fig. 1). We further compile the measurements of all the triads for the day together as shown in Fig. 3(a).

2.2.3 Stage 3: cluster analysis

All the remaining measurements are grouped into clusters (Fig. 3). To properly group the measurements that are detected at different locations and times, we first convert the measured arrival times into distances by multiplying the centroid times with a group velocity of 3 km s^{-1}. The conversion will assure that detection indexes (location and time) are in the same unit (km). We then calculate pairwise distances between all the pairs of the detections based on the 4-D measurement index coordinates (location and time). An agglomerative hierarchical cluster tree is then constructed from the pairwise distances with a cluster cut-off threshold of 3150 km. For example, two triads that are more than 3150 km away or detections of one triad that are 1050 s away are grouped into two different clusters based on the cut-off threshold. This cut-off threshold is implemented to help correlating the detections that are all associated with the same

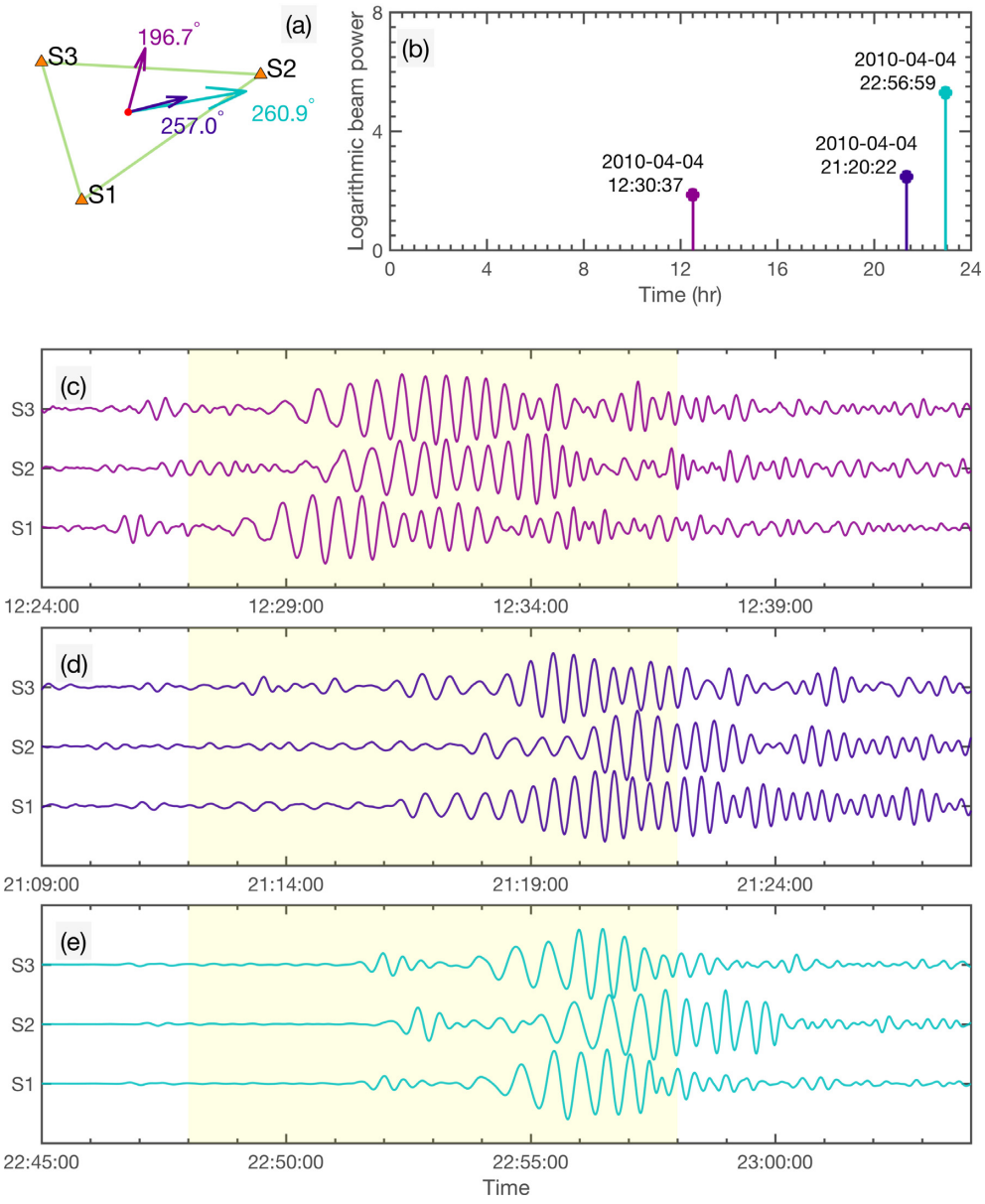

Figure 2. Examples of the coherent signal detections for the example triad in Fig. 1. (a) Propagation directions of the passing surface waves. (b) Centroid time and beam power of the detected signals at the triad. (c)–(e) Bandpass filtered waveforms (20–50 s period) from the three stations (S1–S3). The 600 s time windows used for detecting coherent signals are shown as the yellow shaded regions.

seismic event for the next location stage. The number of groups is decided by the data itself. We empirically remove clusters with less than 75 detections to assure robustness of the coherent signals across the wavefield. As an example, 11 clusters are identified for 2010 April 4 (Figs 3 b and c). Cluster 3 is associated with the 2010 M_w 7.2 El Mayor–Cucapah earthquake in northern Baja California, Mexico (Figs 3 c and 4). The wavefield of cluster 3 is well depicted by the centroid arrival times and the propagation directions (Fig. 4), albeit some inconsistent detections are present in the near-field region (southern California). All the detections of cluster 3 will be used to determine the location and the origin time of the 2010 El-Mayor earthquake. We assume each cluster is only associated with one potential seismic source. It is worth to note that the cut-off threshold distance 3150 s and the minimum detection number 75 to form a cluster are all chosen in an ad hoc fashion, which are dependent on the array configuration and data SNRs.

Instrument response may introduce surface wave phase shifts in the measurements, and can eventually lead to potential biases in location and time. To assess the effects of instrument response, we compare the measured direction of propagations (backazimuth) and arrival times for the same triad at the same time window before and after removing the instrument response (Fig. 5). As an example, instrument response does not cause the measurements of the cluster 3 in 2010 April 4 to vary much (Figs 5 a and c). For all the detections of 2010 April 4, the backazimuth differences are small with a median as 0 (Fig. 5b), while the arrival time differences are non-zero with a median as 14 s (Fig. 5d). The arrival time differences may lead to a 14 s delay on average in determining the seismic event origin times. As we only keep one detection every 300 s, the time delay caused by the instrument response would not bias the final solutions very much.

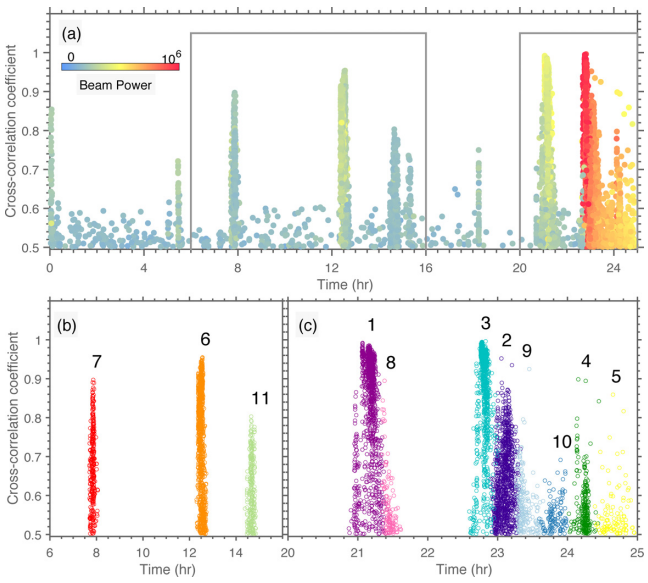

Figure 3. (a) All the detections and (b) and (c) the clustering analysis on 4 April 2010. Vertical axis shows the cross-correlation coefficients. (b) and (c) show clusters identified for the day, which are also windowed out by grey boxes in (a). Cluster 3 associates with the 2010 M_w 7.2 El Mayor–Cucapah earthquake.

2.2.4 Stage 4: seismic sources

Seismic sources are located with the detection clusters in a two-step procedure. First, 10 242 globally evenly spaced gridpoints with 2° distance separation are evaluated for a given cluster (Fig. 7a). At each potential source gridpoint $\underline{x}_i$, the azimuth misfit function $f_a(\underline{x}_i)$ is defined as:

$$f_a(\underline{x}_i) = \frac{1}{N_i}\sum_{j\in G_i}^{N_i}\frac{1}{M_j}|A_j(\underline{x}_i) - A_j^o - \Delta A_j(\underline{x}_i)| \quad (4)$$

where A_j^o is the observed propagation direction at the jth triad, $A_j(\underline{x}_i)$ is the azimuth from the testing grid to the jth triad and $\Delta A_j(\underline{x}_i)$ is an arrival angle anomaly correction for the x_i potential source and jth triad pair. For the testing grid $\underline{x}_i$, the azimuth misfit function $f_a(\underline{x}_i)$ is left as undefined if less than 50 per cent of the detections satisfy $|A_j(\underline{x}_i) - A_j^o - \Delta A_j(\underline{x}_i)| \le A_{\text{fit}}$ in the cluster. A_{fit} is an azimuthal misfit threshold, and is set to be 25°. G_i represents subsets of the detections in the cluster that satisfy $|A_j(\underline{x}_i) - A_j^o - \Delta A_j(\underline{x}_i)| \le A_{\text{fit}}$, N_i is the number of the detections in G_i and M_j is the number of contributing triads within 10° azimuthal bin (36 azimuthal bins). Here, M_j does not change with respect to the potential source location, which is independent from x_i. This weighting parameter (M_j) downweights similar measurement, and helps in improving the azimuthal coverage of the measurements.

Current Rayleigh wave tomographic models cannot explain the observed arrival angle anomalies (e.g. Larson & Ekström 2002; Foster *et al.* 2014). As shown in Fig. 6(a), the observed arrival angle anomalies of cluster 3 in 2010 April 4 range from −20° to 20°, while the modelled arrival angle anomalies are within −5° to 5°. The modelled arrival angle anomalies are calculated with the ray tracing method in Woodhouse & Wong (1986) at 30 mHz (33 s), using the tomographic model presented in Ma *et al.* (2014) expanded at $l = 40$ (∼5°). Therefore, we empirically obtain $\Delta A_j(\underline{x}_i)$ from 1293 globally distributed $M_w \ge 4.5$ earthquakes in the Global Centroid Moment Tensor (GCMT) catalogue, which were recorded by the USArray during 2010 (Figs 6 b and c). For spatially collocated events (within 5 km), the event with the maximum moment magnitude is chosen as the preferred calibration event. Twenty glacial quakes occurred during 2010 in Greenland are also included as calibration events (Veitch & Nettles 2012). For the testing grid x_i, if there are calibration events within 500 km, $\Delta A_j(\underline{x}_i)$ is the measured arrival angle anomaly at the jth triad of the nearest calibration event. Otherwise, $\Delta A_j(\underline{x}_i)$ is set to be zero.

The final misfit function at x_i is defined as

$$f(\underline{x}_i) = \Sigma(G_i)\cdot f_a(\underline{x}_i) \quad (5)$$

where $\Sigma(G_i)$ is the standard deviation obtained from the distribution of $A_j(\underline{x}_i) - A_j^o - \Delta A_j(\underline{x}_i)$ for triads in G_i. $\Sigma(G_i)$ is set as undefined if $\Sigma(G_i)$ exceeds 20° or the absolute value of the mean of the distribution is larger than 15°. For instance, Fig. 7(a) shows the global final misfit map of cluster 3 of 2010 April 4. The gridpoint minimizing eq. (5) is denoted as $\underline{x}_g$.

Within 6° of $\underline{x}_g$, a second step search is performed with denser equally spaced gridpoints (0.25°). With the new set of grids, the gridpoint minimizing eq. (5) is the final optimal seismic source location for the cluster, denoted as $\underline{x}_0$. If none of the gridpoints can explain more than 75 detections ($N_i \le 75$), the algorithm treats the regional search as having failed, and designates the cluster as being inconsistent. The detections that meet $|A_j(\underline{x}_0) - A_j^o - \Delta A_j(\underline{x}_0)| \le A_{\text{fit}}$ are denoted as G_0. The final location estimate of the 2010 El Mayor–Cucapah earthquake is shown in Fig. 7(b). The deviation between this location and its epicentre reported by GCMT is 16.5 km (Ekström *et al.* 2012). With the optimal source location $\underline{x}_0$, we then use centroid arrival times of triads in G_0 to estimate the origin time (t_0) and the average group velocity (v_0) travelling across the whole array:

$$\begin{aligned} &\min \quad \|\underline{t}_o - (t_0 + \tfrac{1}{v_0}\underline{d}_0)\|_1 \\ &\text{subject to} \quad 2.5 \le v_0 \le 6 \end{aligned} \quad (6)$$

where $\underline{t}_o$ is the observed centroid arrival time vector, and $\underline{d}_0$ is the geodesic minimal arc distance vector between $\underline{x}_0$ and the triads. The problem is solved with convex optimization (Vandenberghe & Boyd 1996; Boyd & Vandenberghe 2004; Grant & Boyd 2014, 2008). For cluster 3 of 2010 April 4, the estimated average group velocity is 3.04 km s^{-1} with an origin time difference from the GCMT catalogue of 30 s (Fig. 7c).

With the estimated optimal location and origin time, we can predict the centroid arrival times and propagation directions, assuming geodesic minimal arcs as ray paths (Figs 8 c and d). Taking the 2010 El Mayor–Cucapah earthquake as an example, the predictions agree well with observations (Figs 4 and 8). The arrival time residuals and arrival angle anomalies correlate with some velocity features of surface wave tomography studies (Figs 8 b and e) (Foster *et al.* 2013, 2014; Ekström 2014), and will be explored in Section 4 later.

2.2.5 Stage 5: catalogue quality control

We apply the detection algorithm described above to all the clusters identified during 2010 April, and compile a catalogue with the vetted detections. As each day file is 25 hr long, duplicates would be present when events were recorded around mid-night. To remove these duplicates, any detected events that occurred within 0.1° and 150 s of each other are taken as one event, and only the first detected event is kept in the catalogue. In addition, we conservatively remove events located more than 85° away from USArray centre (39.8°/ − 98.6°) because location accuracy degrades with increasing distance from the array.

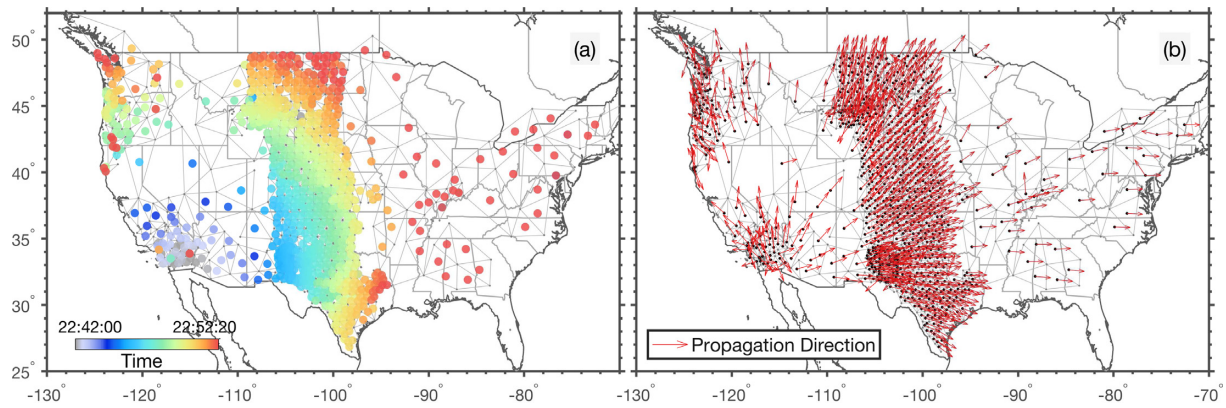

Figure 4. Measurements at all triads. (a) Centroid times and (b) Surface wave propagation directions of cluster 3 in Fig. 3.

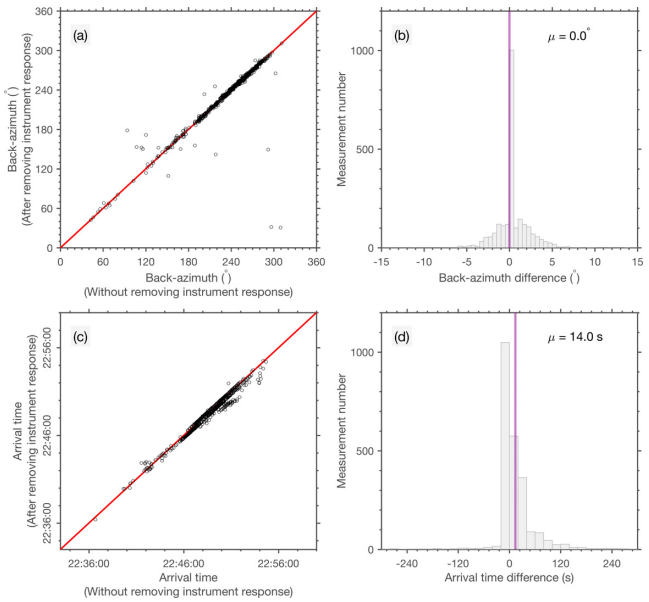

Figure 5. Backazimuth and arrival time compared before and after removing instrument response. (a) and (c) are the measured backazimuth and arrival time of cluster 3 of 2010 April 4 compared between before and after removing instrument response. (b) and (d) are difference distributions of all the measured back-azimuth (median $\mu = 0^\circ$) and arrival time (median $\mu = 14$ s) during 2010 April 4 between before and after removing instrument response.

For detected seismic event, we estimate the location uncertainty by exploring the structure of a suite of grids within a misfit threshold (de Groot-Hedlin & Hedlin 2018). For a given cluster, grids with misfit values $f(\underline{x}_i) \leq 1.25 f(\underline{x}_0)$ are taken as possible source locations. From these grids, a distance covariance matrix is computed, which eigenvectors and eigenvalues are used to define an uncertainty ellipse for the optimal source location (Vasco *et al.* 1993; de Groot-Hedlin & Vernon 1998; de Groot-Hedlin & Hedlin 2018). The optimal source location (x_0) is at the centre of the ellipse, and the uncertainty ellipse minor and major axis lengths are twice the eigenvalues of the covariance matrix, which assures a 95 per cent confidence that the true solution lies within the ellipse. The minor axis length is set to be equal or larger than 30 km because of 0.25° fine grid spacing. As shown in Fig. 7(b), the GCMT location of the 2010 El Mayor–Cucapah earthquake is within the uncertainty ellipse, which minor and major axes are 30 and 45 km. Based on the major ellipse axis length r, we assign the location quality as A ($r \leq 100$ km), B ($100 \leq r \leq 300$ km), or C ($r \geq 300$ km). Details of the detected seismic events are listed in Table S1 in the Supporting Information.

3 RESULTS

We located 271 events within 85° of array centre (39.8°/ − 98.6°) as shown in Fig. 9(a) and Table S1 in the Supporting Information during the month of 2010 April. During this time, the GCMT project (Ekström *et al.* 2012) located 25 events within 75° of the centre of the U.S., with centroid depths that are shallower than 40 km. The minimum moment magnitude of the earthquakes is 4.6 from the GCMT catalogue. By comparing the two catalogues, 23 of the GCMT events were also located with our method (Fig. 9a). If a GCMT event is within 20° and 20 min of our detected location, the detected event is associated with the GCMT event. We missed one event in GCMT that occurred on 2010 April 25 with moment magnitude 4.9, and was located near the Aleutian Islands (GCMT ID: 201004251551A), and another one that occurred on 2010 April 5 (GCMT ID: 201004051114A). These two earthquakes were not detected by enough triads to form clusters at Stage 3, or buried in the coda of a proceeding large earthquakes. In general, the shared events are spatiotemporally close to each other with median spatial deviation of 0.2° and median temporal deviation of 0.8 min (Figs 9 b and 10 a–d). This is because some of these 23 shared events are also included as part of the calibration events.

To understand the location biases introduced by the unknown/inaccurate arrival anomalies, we locate seismic sources with the same set of detections without using calibration events of 2010 April. Within 85° of the array centre, 20 seismic sources are successfully located (Fig. 9c). Another three events were missed, two of which were from the Aleutian Islands and one was from the Caribbean region (Fig. 9c). These three events were located 20° away from the GCMT reported locations. The deviation in both time and location between associated events increase with increasing distance from the array (Figs 10 e–h). Taking the GCMT locations as being exact, our detected events have the largest spatial offsets deviating towards the southwest when the GCMT events are ∼60° away from the array (Fig. 9d).

In addition, we compare our located events with one month of catalogued events in the International Seismological Centre (ISC) bulletin (International Seismological Centre 2013). We compared the located events to ISC events that are shallower than 40 km, larger than M3.5 (any magnitude), and reported by the prime authors. Following the same criteria as used to compare with GCMT catalogue

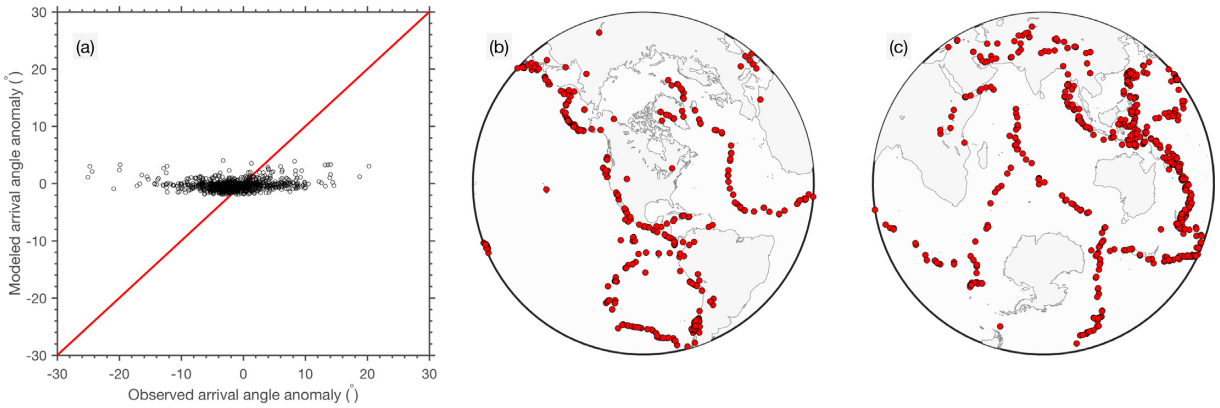

Figure 6. Calibration events and observed and predicted arrival angle anomalies for cluster 3 of 2010 April 4. (a) Observed arrival angle anomaly versus predicted arrival angle anomaly from the tomographic model in Ma *et al.* (2014) with the ray tracing approach form Woodhouse & Wong (1986) at 30 mHz (33 s). (b) and (c) are all the calibration events used during the location step.

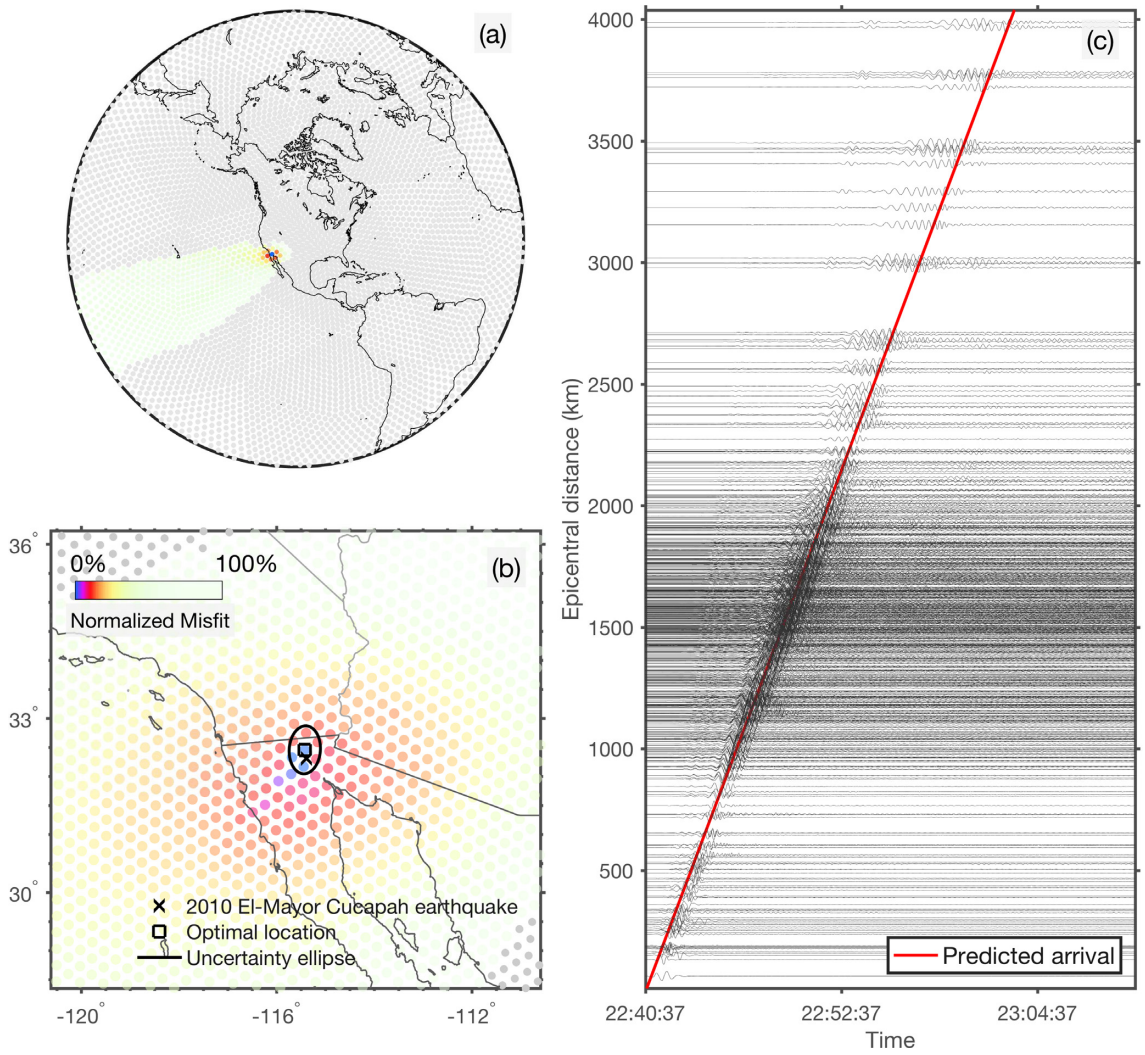

Figure 7. Detection of the 2010 M_w 7.2 El Mayor–Cucapah earthquake. (a) Global misfit map. The misfit is normalized to the maximum value. (b) Regional misfit map. The misfit is normalized, and both (a) and (b) share the same colour bar. Grey dots indicate undefined values. (c) Bandpass filtered waveforms aligned with the detected epicentre and the origin time. The predicted arrival from the estimated seismic source location, origin time and average group velocity is shown as the red line.

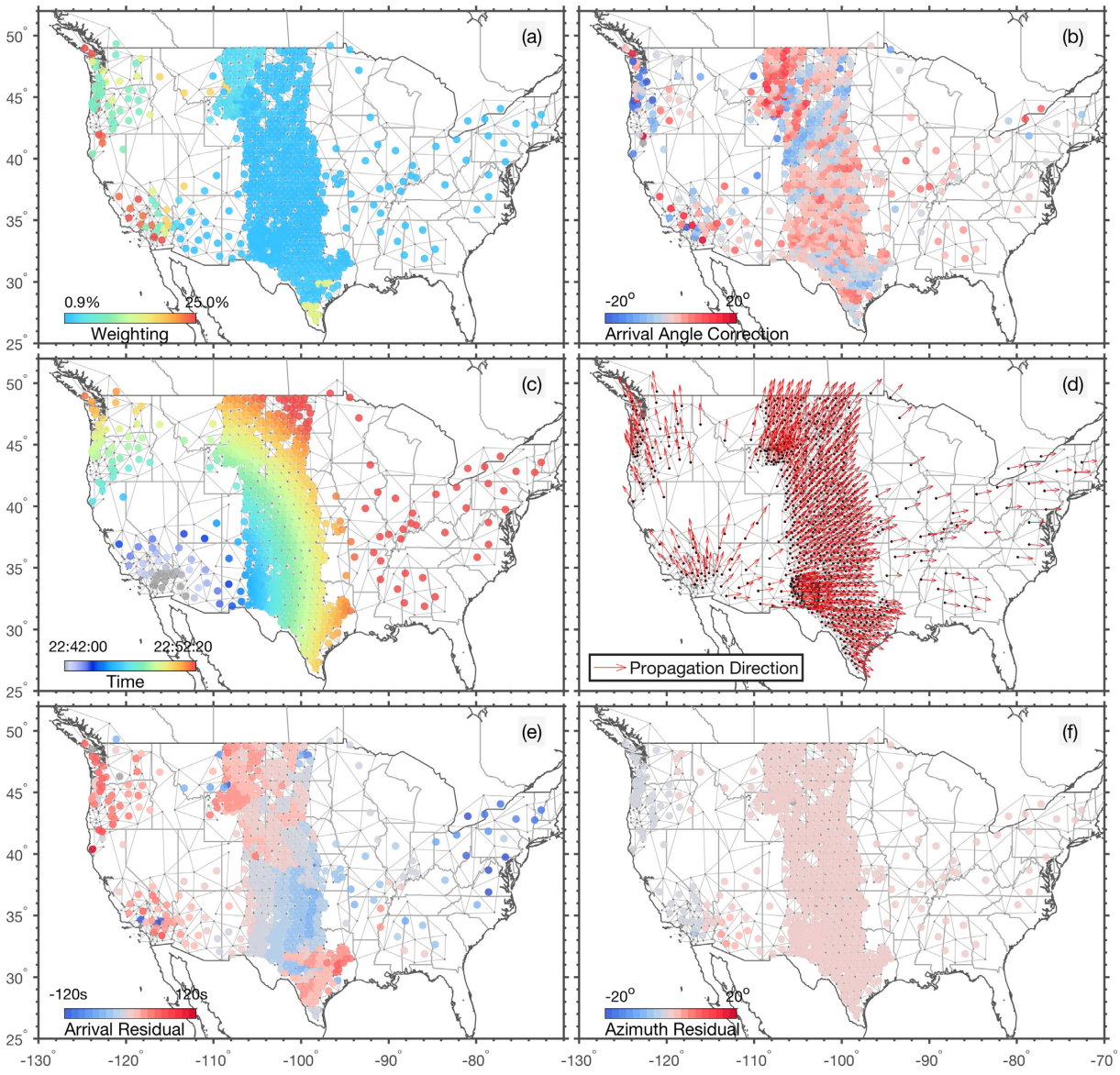

Figure 8. Weighting, arrival angle anomaly corrections, predicted arrival times and propagation directions at each triad from the detected seismic source in Fig. 7. Only triads used for the grid search are plotted. (a) Weighting. (b) Arrival angle anomaly corrections for optimal location. (c) Centroid times. (d) Surface wave propagation directions calculated assuming great circle propagation (without arrival angle anomaly corrections). (e) and (f) Arrival time residuals and azimuthal residuals between the prediction and the measurements in Fig. 4 after applying the arrival angle anomaly corrections.

(within 20° and 20 min), 131 events are listed in both the ISC catalogue and our catalogue. In addition, there are another 140 events in our catalogue do not associate with the ISC catalogue. Some of these events could have been located at erroneous locations, and some of these 140 events could be abnormal events that are deficient in high-frequency seismic radiation as ISC uses short-period body wave phase picks and globally distributed stations to locate earthquakes.

4 DISCUSSIONS

4.1 Earthquake case study: the 2010 M_w 7.2 El Mayor–Cucapah earthquake sequence

We tested the algorithm with the 2010 M_w 7.2 El Mayor–Cucapah earthquake sequence. The 2010 El Mayor–Cucapah earthquake sequence involves two foreshock sequences, one in 2010 March, and another 24 hr before the main shock, including an M4.4 foreshock (Hauksson *et al.* 2011). The 2010 El Mayor–Cucapah main shock ruptured more than four faults, producing over 10 000 aftershocks and causing an aftershock zone extending over 120 km (Hauksson *et al.* 2011; Wei *et al.* 2011). Over 88 $M \geq 4$ aftershocks occurred from April 4 to the end of the month (Hauksson *et al.* 2011). Most were not listed in the GCMT catalogue because of their relatively low magnitudes and because they were deficient in low-frequency surface waves. For instance, the M_w 4.34 foreshock was not listed in the GCMT catalogue, although it excited intermediate-period surface waves that were observed across the whole U.S. continent (Fig. 11a). To evaluate our detection capability, we compared our results with the SCSN earthquake catalogue during 2010 April (Hauksson *et al.* 2011). If the time offset between a detected event and an SCSN reported aftershock is less than 15 min and the distance offset is less than 2°, the detected event is associated with the

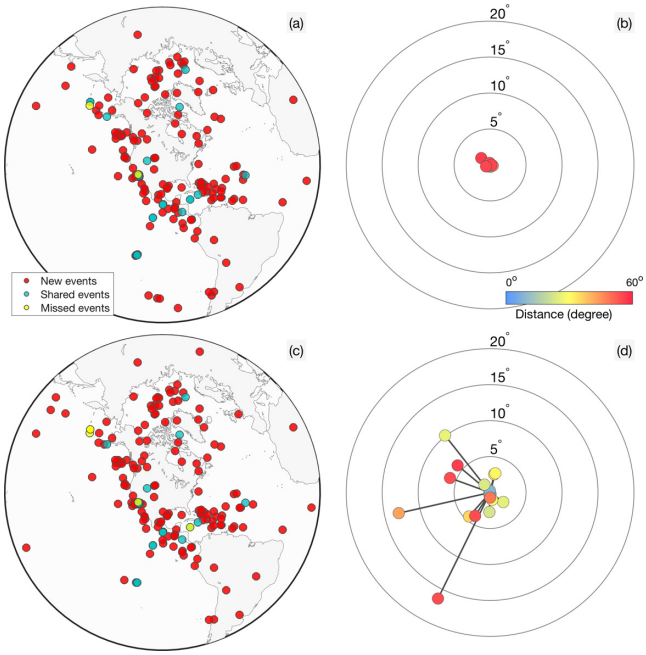

Figure 9. Located events within 85° of the array centre (39.8° / − 98.6°) during 2010 April. (a) Map view of 23 events that were detected by both our approach and the GCMT, two event that were reported in GCMT but missed in our catalogue, and 246 events that were not registered at GCMT. Some of the 23 shared events overlap with each other in the figure when they are spatially close. The events are located with corrections from all the calibration events. (b) Spatial deviations from the GCMT locations of the 23 events in common. The GCMT locations are in the centre, the radii show the deviation distance, and the azimuths indicate the deviation direction. The dot colours show the distances between the centre of each array and the GCMT reported locations. (c)/(d) are similar to (a)/(b), but without calibration events which occurred in 2010 April.

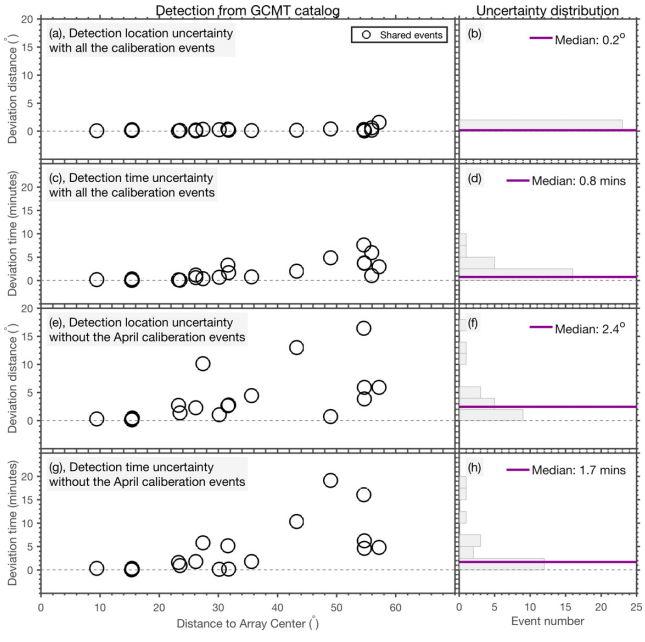

Figure 10. Shared events within 75° of array centre (39.8° / − 98.6°) during 2010 April. (a) Diagram of location deviations of the shared events versus event distances to array centre using all the calibration events. (b) Histogram of the deviation distances in (a). (c) Diagram of time deviations of the shared events using all the calibration events. The legends are similar to those in (a). (d) Histogram of the deviation times in (c). (e)–(h) are similar to (a)–(d), but without calibrations events occurred during 2010 April.

SCSN reported aftershock (Fig. 12). In addition to the M_w 4.34 foreshock and the M_w 7.2 main shock, there are another 54 detections in common with the SCSN reported aftershocks, including 36 $M \geq 4$ aftershocks and 18 $3.4 \leq M \leq 4$ aftershocks. The results show the majority of the $M \geq 4$ were missed in our catalogue (52 out of 88). The incompleteness of the catalogue may have resulted from our relatively strict data processing criteria. For instance, we only allow one measurement per 300 s, which would eliminate closely spaced events and only keep the one with maximum surface wave amplitude in the time window. We also require that each cluster has a minimum 75 triads, which could have missed events that did not efficiently excite an intermediate-period surface wavefield. In addition, successful detections have to explain more than 75 back-azimuth measurements. Furthermore, each cluster is assumed to associate with only one event, which may not be the case for the active aftershock sequence. On the other hand, the majority of the $M \geq 4$ aftershocks occurred soon after the M_w 7.2 main shock, and were buried in the main shock coda (Hauksson *et al.* 2011). The noisy data also caused challenges of detections with our proposed method.

The smallest aftershock detected by our algorithm is a M_w 3.42 earthquake that occurred on April 5 (Fig. 11b). Although only a fraction of the aftershocks were detected by our algorithm, the spatial deviation of the events in common are small with a median of 0.2° (insert of Fig. 12a). With our catalogue, the aftershock temporal decay roughly follows Omori's law, with parameters modified from Hauksson *et al.* (2011) (Fig. 12d, $p = 1.01$, $c = 1.5$ and $\kappa = 20$). The median of the logarithmic beam powers from all the triads of a given detection correlates with the earthquake moment magnitude (Fig. 12c). This is likely because the triads that detect all of these events are roughly the same, and beam powers are proxies of surface wave amplitudes as most of the instruments are the same. In addition, all of these aftershocks have similar focal mechanisms, which leads to similar radiation patterns. Therefore, the correlation is not surprising. However, it might not be the case for sources located elsewhere. The results suggest that our catalogue is not complete for a given magnitude, for example M_w 4. However, the successful detections of the ∼M3.5 earthquakes highlight that the algorithm is capable of detecting small events with high noise level data (Fig. 11b).

4.2 Non-earthquake case study: a glacial earthquake and a landslide in Greenland

One advantage of our algorithm is that it uses simple array processing to locate seismic sources without requiring seismic phase picks. Therefore, no *a priori* assumptions about the source type are needed, which enables the algorithm to detect non-earthquake sources such as glacial quakes. This new class of moderate size quakes was first identified and located by global array techniques (Ekström *et al.* 2003, 2006; Tsai & Ekström 2007; Nettles & Ekström 2010; Chen *et al.* 2011). These glacial quakes are generated by ice motions, which often last 15–30 times longer than tectonic earthquakes, and are depleted at high frequencies (Ekström *et al.* 2003, 2006; Tsai & Ekström 2007; Joughin *et al.* 2008; Nettles *et al.* 2008; de Juan *et al.* 2010; Nettles & Ekström 2010; Chen *et al.* 2011; Murray *et al.* 2015a,b). Therefore, surface wave detectors are better suited to locating them, as traditional methods rely on high-frequency body wave phase picks. The majority of the glacial quakes that are located in Greenland are associated with large ice-calving events (Tsai & Ekström 2007; Veitch & Nettles 2012; Olsen & Nettles 2017).

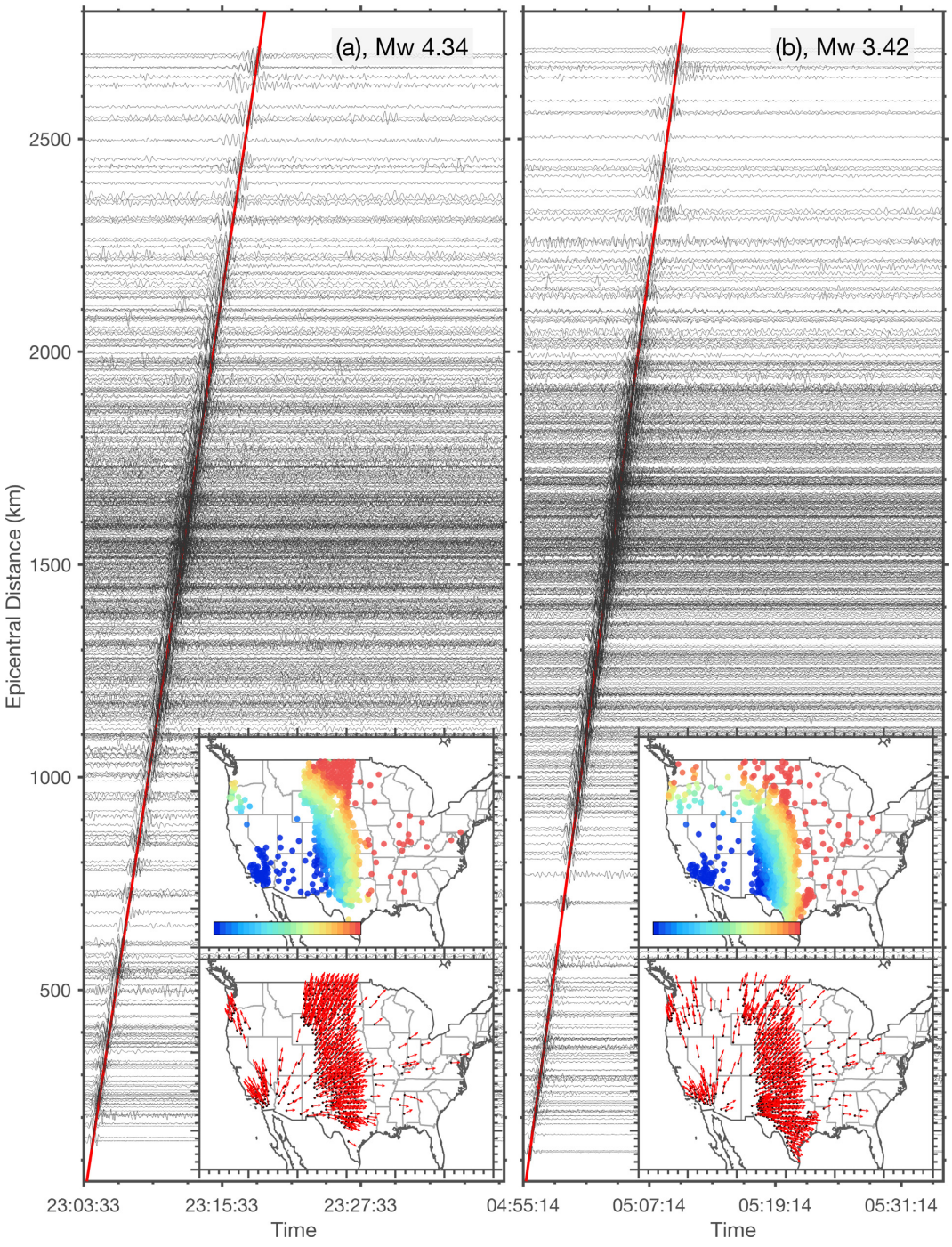

Figure 11. Measurements at all triads, predicted arrivals from detected seismic sources and the aligned waveforms. (a) Foreshock of the 2010 M_w 7.2 El Mayor–Cucapah earthquake. (b) The smallest detected aftershock during the month. The legends are similar to those in Figs 4 and 7.

During 2010 April, our algorithm detected and located two events (2010 April 27) in Greenland. One event (03:25:25) located at the Daugaard Jensen region is likely to be a glacial earthquake (Fig. 13a). This event is detected by 702 triads, and the associated waveforms of the seismic source are anomalously long and are clearly high-frequency depleted (Fig. 13b). When single traces have a high noise level, the emergent wave trains can be challenging to identify without a large set of stations (Fig. 13b). This possible glacial quake was not reported in Veitch & Nettles (2012), which might be because of its poor SNR for globally distributed stations. The other event (18:40:40) is located interior of Greenland without clear association with any of the previous identified regions frequently hosting glacial quakes. During 2010 April, one glacial quake (2010 April 14) occurred in the Jakobshavn Isbrae region with an equivalent magnitude of 4.7 (Veitch & Nettles 2012). Signals associated with this glacial quake cannot be easily identified with the given array configuration, and our algorithm did not form a detection cluster for this event.

We performed our algorithm to previously reported glacial quakes occurred in Greenland to understand the location uncertainties of

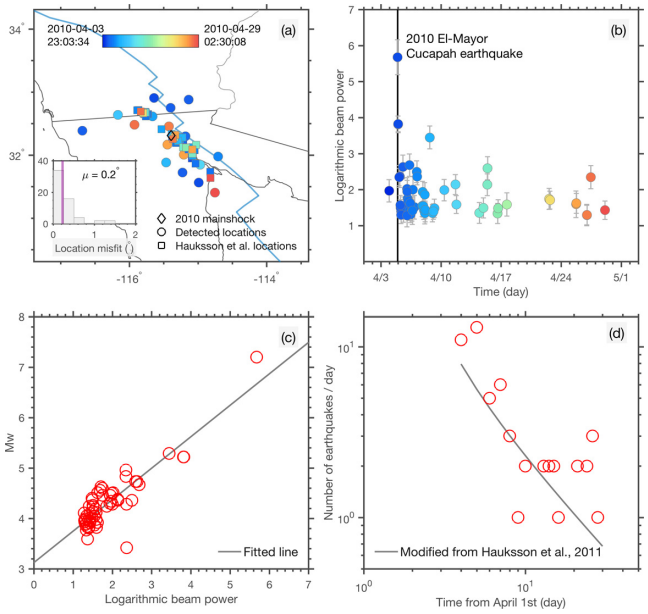

Figure 12. The 2010 M_w 7.2 El Mayor–Cucapah earthquake sequence. (a) Map view of the detected events and their locations reported in Hauksson *et al.* (2011). The insert is the histogram showing the deviations between the detected locations and those reported in Hauksson *et al.* (2011). (b) Beam powers of the detected events versus time. The beam power of every event is the median of the measurements from all the triads that recorded the event. The error bar shows one standard deviation of the beam power. (c) Logarithmic beam power versus moment magnitude (M_w) of the events. The logarithmic beam powers are the same as in (b), and the moment magnitudes are from Hauksson *et al.* (2011). The fitted line has a scaling factor of 0.68 and a constant of 3.04. (d) The decay of the detected El Mayor–Cucapah aftershocks showing number of events per day versus time; the fitted line is modified with parameters from Hauksson *et al.* (2011).

the detected seismic events in the region. Thirty glacial quakes occurred in Greenland during 2010 (Veitch & Nettles 2012). Twenty of these events were detected with the given array configuration and formed clusters (Fig. 13a). With arrival angle anomaly corrections, these events are well located into six regions with median location deviation less than 0.2° (insert of Fig. 13). This validation test suggests that our approach can resolve glacial quakes in Greenland with high precision (0.2°) when there are nearby calibration events. We further test the location accuracy by applying the approach without arrival angle anomaly corrections (insert of Fig. 13). For the same set of events, they are located within 5° on average to the locations reported in Veitch & Nettles (2012). These results suggest our method is complementary to existing methods and can help in building a more complete glacial quake catalogue to investigate seismogenic glacial movements.

Landslides are quakes of another class that can efficiently excite surface waves (e.g. Iverson *et al.* 2015; Gualtieri & Ekström 2016, 2018). Landslides which cause large mass movements in a short period of time can have devastating effect (e.g. Hibert *et al.* 2015; Gualtieri & Ekström 2016, 2018). Despite their hazard relevance, it has been difficult to set up experiments to monitor the phenomena because of an incomplete understanding of their physical mechanisms and triggering patterns. Seismic monitoring has been one of the major tools to investigate such processes because landslides can generate seismic signals with a broad band of frequency content (e.g. Hibert *et al.* 2014, 2015; Gualtieri & Ekström 2016, 2018; Poli 2017). For instance, a large destructive landslide occurred near Nuugaatsiaq in northwestern Greenland on 2017 June 17, and generated clear seismic signals that were observed across the continental U.S., including Alaska (Fig. 14). This landslide also induced a local tsunami, which also generated long-period seismic signals with dominant period around 200 s (Bessette-Kirton *et al.* 2017). Waveform analysis from United States Geological Survey suggests that the Nuugaatsiaq landslide released an equivalent amount of energy as an M_s 4.8 earthquake (Bessette-Kirton *et al.* 2017). We use this landslide as a benchmark to test the utility of our algorithm for detecting such anomalous seismic sources. We have successfully detected and located the landslide, with our epicentre deviating 75 km from the location resolved from satellite imagery (Fig. 14). Similar to the procedure described in Section 2.2, we used multiple arrays to locate the landslide, including stations in southern California, eastern U.S., Alaska and Greenland. No arrival angle anomaly corrections are applied here. The effective combination of these multiple arrays provides a good azimuthal coverage, leading to the small location error at the large distance range when compared to only having stations in the central U.S. (Fig. 13a). The results also suggest that our algorithm can easily accommodate an arbitrary number of large arrays with various array geometries. The successful detections of the glacial quakes and the landslide demonstrate the potential of using our algorithm to study anomalous seismic sources.

4.3 Uncertainties and limitations

Understanding the quality of the detected seismic source parameters is essential for use of the catalogue in probing the physical mechanisms of these sources. As shown in Figs 9 and 10, resolutions of the estimated source location and origin time depend on where the source is located. The algorithm performs the best for detecting events in the region of Gulf of California. For instance, the small deviations of the 2010 El Mayor–Cucapah earthquake sequence (~25 km). However, for 2010 April, the increasing deviations from the GCMT reported locations and time of events beyond 30° of the array degrade the catalogue. This is because the algorithm is completely data driven, and the resolution is heavily influenced by the array geometry. For 2010 April, the array mostly covered a narrow longitude band, which was suboptimal for events that occurred along the Aleutian Islands and the South American subduction zone. This can be partially mitigated by applying arrival angle anomaly corrections from calibration events. However, such calibration cannot be applied uniformly in space or time due to the eastward evolution of the USArray. As a consequence, our method does not have a uniform resolution for events at all azimuths for the given network. Similar issues related to array geometry have been reported in de Groot-Hedlin & Hedlin (2015). This limitation can be potentially mitigated with more stations. In the example shown in Fig. 14, the spatial resolution is within 1° for the Nuugaatsiaq landslide even without arrival angle anomaly corrections, which is 40° away from the array centre (39.8°/ − 98.6°). The improved resolution comes from incorporating both stations in the continental U.S. and the USArray in Alaska, providing an extended azimuthal coverage.

Path deviations off great circles (geodesic minimal arc) is the main source of uncertainty. In our approach, surface waves are assumed to propagate along the minor-arc great circles. However, surface wavefield can have great complexities because of lateral refraction, scattering and multipathing (e.g. Bungum & Capon 1974; Laske 1995; Laske & Masters 1996, 1998; Ji *et al.* 2005; Larson

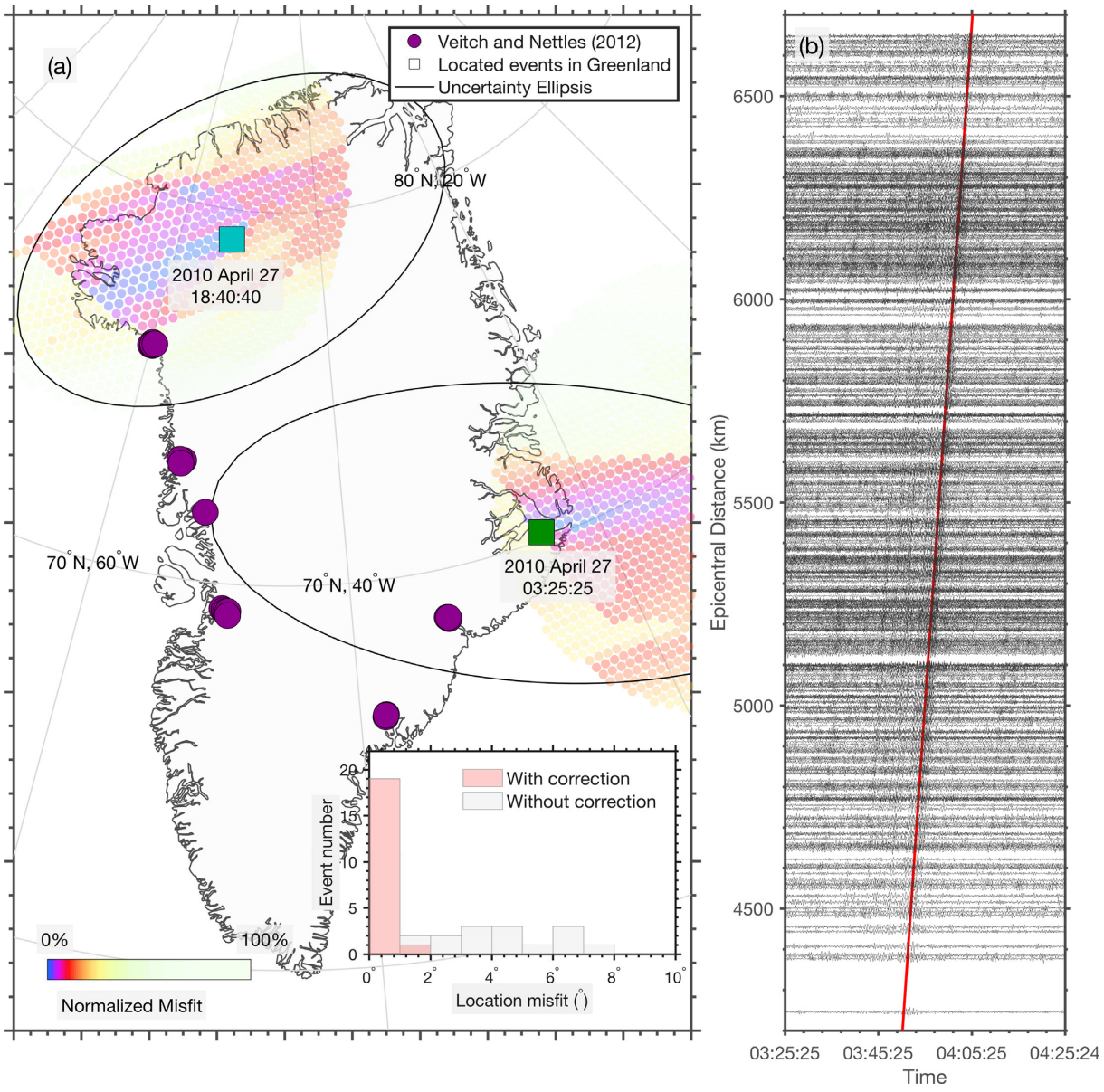

Figure 13. Detections of two seismic events in Greenland. (a) Twenty glacial quakes used for the validation tests and two seismic events located in Greenland. Insert: location deviation distribution for the 20 glacial quakes with and without arrival angle anomaly corrections. (b) Aligned waveforms for the potential glacial earthquake occurred during 2010 April 27 03:25:25. The legends are similar to those in Figs 4 and 7.

& Ekström 2002; Foster *et al.* 2014; Pedersen *et al.* 2015; Doran & Laske 2017). These complexities can result in deviations from the great circles, causing arrival angle anomalies. The arrival angle anomalies are significant for paths along the boundaries between oceans and continents (e.g. Ma *et al.* 2015). For instance, tomographic velocity models can suffer from artefacts when including ray paths that travel along the northern rim of the Pacific Ocean (e.g. Dalton & Ekström 2006; Ruan & Zhou 2012; Ma *et al.* 2015; Bao *et al.* 2016). Previous studies have reported arrival angle anomalies up to $\sim\pm30^{\circ}$ for extreme cases at the 20–50 s period (Laske 1995; Laske & Masters 1996; Tanimoto & Prindle 2007). The arrival angle anomalies across the USArray are within $\pm10^{\circ}$ at 50 s (Foster *et al.* 2014). The backazimuth observed at each triad in our approach describes the directionality of the incoming wave, which may deviate from great-circle path because of the wavefield complexities between the source and the triad. Once the assumptions are not met, and the array is not large enough to average out the anomalies, location bias would be introduced when using the backazimuth to determine the seismic source location. When most of the stations are in the central U.S., effects of such biases are the strongest for events along the Aleutian Islands, East Pacific Rise and the South American subduction zone as observed in previous studies and in our analysis (e.g. Laske 1995; Laske & Masters 1996; Tanimoto & Prindle 2007; Ma *et al.* 2015). On the contrary, influences are marginal for surface waves propagating along the east–west direction. Therefore, the global 3-D velocity complexities also cause non-uniform catalogue resolution at different directions. As shown in Fig. 6(a), implementing surface wave ray tracing with current velocity models is inadequate to model the observed arrival angle anomalies (e.g. Woodhouse & Wong 1986; Ma *et al.* 2015). This challenge has also been noted in Larson & Ekström (2002). Instead, we apply an arrival angle anomaly correction term ($\Delta A_j(\underline{x}_i)$) to empirically correct the

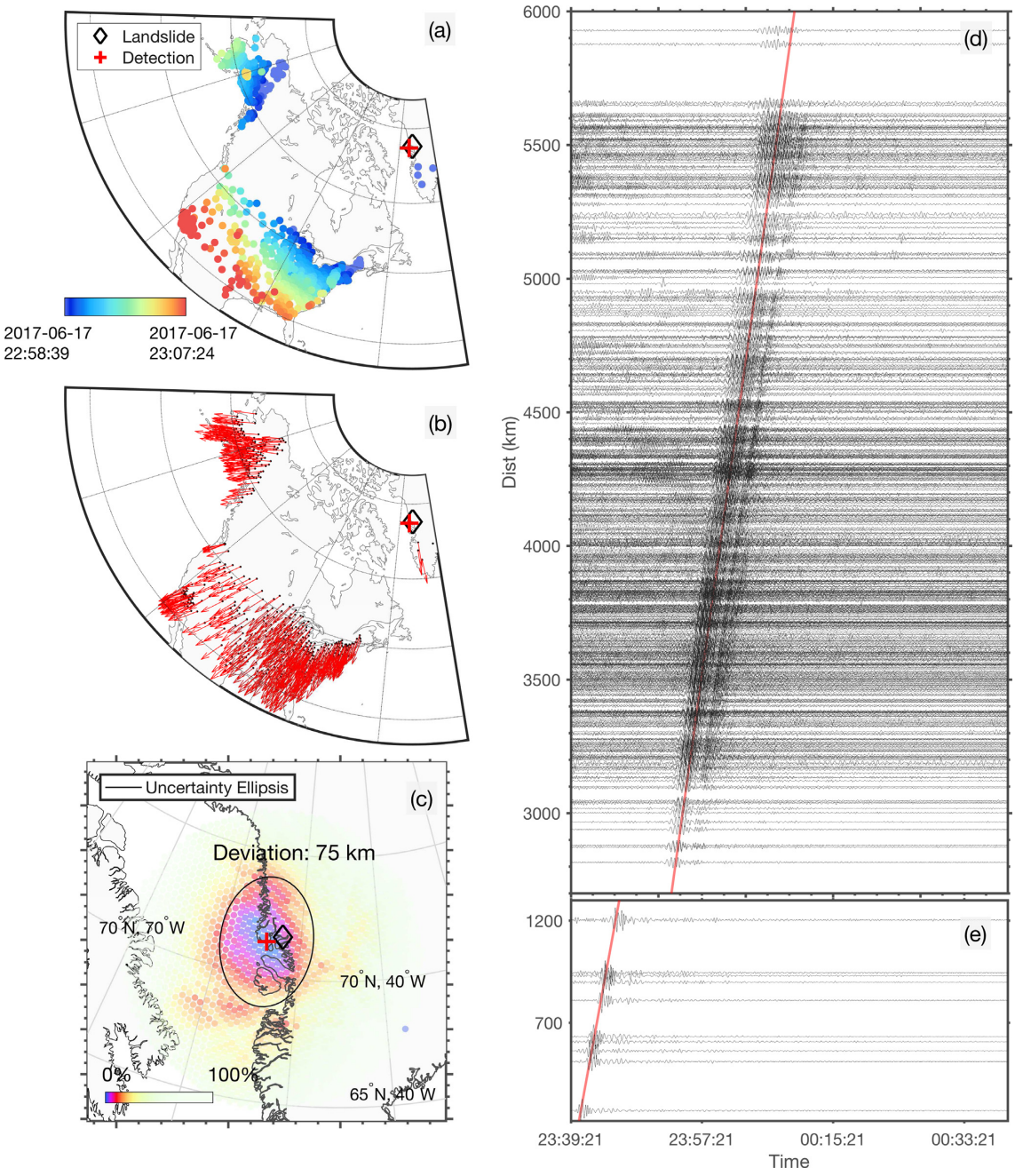

Figure 14. Detection of the 2017 Nuugaatsiaq landslide in Greenland. The legends are similar to those in Figs 4 and 7. The black diamond is the location resolved from satellite imagery, and the red cross is the optimal location from our study.

deviation for the source–receiver pair $\underline{x}_i$) and the jth triad. This can effectively improve the resolution when calibration events are available (Figs 10 a–e). However, as discussed above, non-continuous occurrence and non-uniform distribution of the calibration events limit the corrections only to regions with active seismicity. In general, incorporating globally distributed multiple large aperture arrays can potentially help in improving the catalogue resolution (Fig. 14). On the positive side, arrival angle anomalies obtained by our method while locating seismic sources can be used to constrain velocity models (Fig. 8). Traditionally, arrival angle anomaly measurements often involve cross-correlations with synthetic waveforms, and can be challenging to obtain (Laske 1995; Larson & Ekström 2002; Foster *et al.* 2014). Our results indicate new directions to efficiently make such measurements. In addition, the obtained arrival time residuals and resolved local phase velocities within each triad may also help in constraining local tomographic models (Fig. 8).

We assume a planer signal travelling through the triads. In reality, scattering and multipathing can be significant causing the wavefield to be distorted from plane waves for regional arrays (e.g. Friederich & Wielandt 1995; Forsyth & Li 2005). Luckily, the majority of the triads formed with the USArray are spatially compact with a median side length as 76 km (Fig. 1). This array configuration helps mitigating the possible bias introduced by the complex incoming surface waves. In addition, arrival angle anomaly corrections help in removing the unknown biases, and the azimuthal

deviation threshold ($A_{\rm fit}$) assures that measurements with poor qualities are removed during the location step. The large set of detections from the USArray can provide a wide azimuthal coverage for most of the seismically active regions. This major advantage helps neutralizing possible biases from a small set of triads. Finally, the uncertainty ellipses of the resolved seismic sources provide a quantitative evaluation of the location qualities (Table S1, Supporting Information).

Our proposed method cannot resolve the seismic source depths nor their magnitudes. Surface wave amplitudes are sensitive to the source radiation patterns, ray path structures and the source depths (Gilbert 1971; Dziewoński *et al.* 1981; Ekström *et al.* 2012). By performing waveform modelling of multiple phases, GCMT is able to resolve the event depths (Ekström *et al.* 2012). In our formulation, only the propagation directions of surface waves passing each triad are used to locate the events. Therefore, we cannot resolve the source depths. To resolve magnitudes of seismic sources, seismic wave phase amplitudes (e.g. surface wave) need to be modelled. As our approach does not assume a velocity model and does not use the amplitude information, magnitudes remain elusive to resolve. For special cases like the 2010 $M_{\rm w}$ 7.2 El Mayor–Cucapah earthquake sequence, triad beam powers can be used to empirically calibrate for event magnitudes as they are correlated with the earthquake surface wave amplitudes and the array was unchanged for that period of time. Such calibrations can be performed for a given region with a fixed set of stations, but are limited when the USArray moves from the West Coast to East Coast. These disadvantages stem from not using the velocity structure information. However, the ability to detect events without a known velocity model is one of the major advantages of the approach. Our formulation enables it to be used to detect events at regions with unknown complex structures.

The proposed method of the study is an empirical approach to locate seismic sources with surface waves, which involves a set of experiential parameters. For instance, the frequency band of the surface waves, the criteria of Delaunay triangulation, time window length, the group velocity (3 km s^{-1}) used for converting centroid arrival time to distance, cross-correlation coefficient cut-off threshold and the weighting of local dense arrays are empirically chosen for this study. As for any detection-location approaches, trade-offs exist between detection threshold and location accuracy. The parameters used in this study can serve as initial values, but should be adjusted to accommodate different arrays or targeted scientific problems. For abnormally long wave trains, such as the surface wave packages in Fig. 13(b), the resolved centroid time at each triad may experience cycle skipping. This is because we only allow one measurement every five minutes. We have experimented with a few sets of different parameters, and the current choice are only tested with the given array geometry. The algorithm with this set of parameters can serve as a useful reconnaissance tool screening through large volumes of data to identify regions/events with anomalous features. In its current form, it is intend to be the base for a more focused method specifically designed to uncover events of interest. The 25° threshold of arrival angle misfit is based on the arrival angle anomalies reported in previous studies (Laske 1995; Laske & Masters 1996; Tanimoto & Prindle 2007). The use of the ℓ_1 norm in eq. (4) aims to reduce the influence of outliers, assuming that the misfit follows the Laplace distribution. Systematic evaluation of the location/timing uncertainties will involve multiple years of observations, and the outcomes will be discussed in a follow-up study with comprehensive analysis with data from the whole deployment of the USArray.

5 CONCLUSIONS

We have developed a new method to detect seismic sources with intermediate-period surface waves recorded by large aperture arrays. By dividing large arrays into non-overlapping triads, surface wave coherence between adjacent stations is exploited to determine the propagation direction (backazimuth) of the incoming waves at many discrete points. With the resolved backazimuth, a grid search is then performed to locate the seismic sources. The method does not require *a priori* knowledge of the seismic source types nor a velocity model to locate the sources. Combinations of intermediate-period band surface waves and dense arrays guarantee coherent waveforms, simplify cluster identification, promote high spatiotemporal resolution and reduce false detections. Presumably, the method can also be applied to globally widespread triads of sensors when the detections are properly grouped together. We demonstrate the practicality of the approach by applying it to U.S. continental arrays of 2010 April. The method located the foreshock, the main shock and 56 aftershocks of the 2010 $M_{\rm w}$ 7.2 El Mayor–Cucapah earthquake sequence with median location uncertainty of ∼25 km and event magnitude as low as $M_{\rm w}$ 3.4. We successfully detected a potential glacial quake in Greenland during the month which was not reported in previous global array studies. The method can also locate exotic seismic sources such as landslides in addition to glacial quakes. These results indicate the potential to use the approach to identify new classes of earthquakes. Furthermore, the proposed method can potentially aid in providing extra measurements for global and regional tomographic studies such as arrival time residuals and arrival angle anomalies. By dividing a large aperture array into triads, our approach has the flexibility to incorporate multiple spatially disconnected large arrays. The method is easy to implement, very general, and can potentially be applied in real time.

ACKNOWLEDGEMENTS

We thank the editor Dr. Frederik Simons, Dr. Steve Gibbons and the other reviewers for their constructive suggestions, which have led to improvements in our paper. We thank A. Doran, J. McGuire and P. Shearer for helpful discussions. The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681. The earthquake catalogue was downloaded from the Global Centroid Moment Tensor (GCMT, globalcmt.org, last accessed 2018 July 12) project (Ekström *et al.* 2012) and the International Seismological Center (ISC, isc.ac.uk, last accessed 2018 July 12) (International Seismological Centre 2013). WF is currently supported by the Postdoctoral Scholar Program at the Woods Hole Oceanographic Institution, with funding provided by the Weston Howland Jr. Postdoctoral Scholarship. This work was supported by National Science Foundation grant EAR-1358520 at Scripps Institution of Oceanography, UC San Diego. The AELUMA code used to detect and locate either infrasound, body wave, or surface wave events can be obtained on request through https://ds.iris.edu/ds/products/infrasound-aeluma/#request (last accessed 2018 July 12). All the processed data are available upon request to the authors.

REFERENCES

Bao, X., Dalton, C.A. & Ritsema, J., 2016. Effects of elastic focusing on global models of Rayleigh wave attenuation, *Geophys. J. Int.,* **207**(2), 1062–1079.

Bessette-Kirton, E., Allstadt, K., Pursley, J. & Godt, J., 2017. *Preliminary analysis of satellite imagery and seismic observations of the Nuugaatsiaq landslide and tsunami,* Greenland, USGS, Washington, DC. Available at: https://landslides.usgs.gov/research/featured/2017-nuugaatsiaq/, (Last accessed 2018 July 12).

Boyd, S.P. & Vandenberghe, L., 2004. *Convex Optimization,* Cambridge University Press.

Bungum, H. & Capon, J., 1974. Coda pattern and multipath propagation of Rayleigh waves at NORSAR, *Phys. Earth planet. Inter.,* **9**(2), 111–127.

Busby, R., Vernon, F., Newman, R. & Astiz, L., 2006. EarthScope's Transportable Array: Advancing eastward, *EOS, Trans. Am. geophys. Un.,* **87**(52), Abstract ID: U41B-0820.

Cansi, Y., 1995. An automatic seismic event processing for detection and location: The PMCC method, *Geophys. Res. Lett.,* **22**(9), 1021–1024.

Cattania, C., McGuire, J.J. & Collins, J.A., 2017. Dynamic triggering and earthquake swarms on East Pacific Rise transform faults, *Geophys. Res. Lett.,* **44**(2), 702–710.

Chen, X., Shearer, P.M., Walter, F. & Fricker, H.A., 2011. Seventeen Antarctic seismic events detected by global surface waves and a possible link to calving events from satellite images, *J. geophys. Res.,* **116**(B6), doi:10.1029/2011JB008262.

Dalton, C.A. & Ekström, G., 2006. Global models of surface wave attenuation, *J. geophys. Res.,* **111**(B5), doi:10.1029/2005JB003997.

de Groot-Hedlin, C.D. & Hedlin, M.A., 2015. A method for detecting and locating geophysical events using groups of arrays, *Geophys. J. Int.,* **203**(2), 960–971.

de Groot-Hedlin, C.D. & Hedlin, M.A., 2018. A new automated approach to detecting and locating seismic events using data from a large network, *Bull. seism. Soc. Am.* , **108,** 4, 2032–2045.

de Groot-Hedlin, C.D. & Vernon, F.L., 1998. An evolutionary programming method for estimating layered velocity structure, *Bull. seism. Soc. Am.,* **88**(4), 1023–1035.

de Groot-Hedlin, C.D., Hedlin, M.A. & Walker, K.T., 2014. Detection of gravity waves across the USArray: a case study, *Earth planet. Sci. Lett.,* **402,** 346–352.

de Groot-Hedlin, C.D., Hedlin, M.A.H., Hoffmann, L., Alexander, M.J. & Stephan, C.C., 2017. Relationships between gravity waves observed at Earth's surface and in the stratosphere over the central and eastern United States, *J. geophys. Res.,* **122**(21), 11 482–11 498.

de Juan, J. *et al.*, 2010. Sudden increase in tidal response linked to calving and acceleration at a large Greenland outlet glacier, *Geophys. Res. Lett.,* **37**(12), doi:10.1029/2010GL043289.

Doran, A.K. & Laske, G., 2017. Ocean-bottom seismometer instrument orientations via automated Rayleigh-wave arrival-angle measurements, *Bull. seism. Soc. Am.,* **107**(2), 691–708.

Douglas, A., 1967. Joint epicentre determination, *Nature,* **215,** 47–48.

Dziewoński, A.M., Chou, T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86**(B4), 2825–2852.

Ekström, G., 2014. Love and Rayleigh phase-velocity maps, 5–40 s, of the Western and Central USA from USArray data, *Earth planet. Sci. Lett.,* **402,** 42–49.

Ekström, G., Nettles, M. & Abers, G.A., 2003. Glacial earthquakes, *Science,* **302**(5645), 622–624.

Ekström, G., Nettles, M. & Tsai, V.C., 2006. Seasonality and increasing frequency of Greenland glacial earthquakes, *Science,* **311**(5768), 1756–1758.

Ekström, G., Nettles, M. & Dziewoński, A., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200,** 1–9.

Ekström, G., 2006. Global detection and location of seismic sources by using surface waves, *Bull. seism. Soc. Am.,* **96**(4A), 1201–1212.

Fan, W. & Shearer, P.M., 2016. Local near instantaneously dynamically triggered aftershocks of large earthquakes, *Science,* **353**(6304), 1133–1136.

Forsyth, D.W. & Li, A., 2005. Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference, *Seism. Earth: Array Anal. Broadband Seismog.,* **157,** 81–97.

Foster, A., Ekström, G. & Nettles, M., 2013. Surface wave phase velocities of the Western United States from a two-station method, *Geophys. J. Int.,* **196**(2), 1189–1206.

Foster, A., Ekström, G. & Hjörleifsdóttir, V., 2014. Arrival-angle anomalies across the USArray Transportable Array, *Earth planet. Sci. Lett.,* **402,** 58–68.

Friederich, W. & Wielandt, E., 1995. Interpretation of seismic surface waves in regional networks: joint estimation of wavefield geometry and local phase velocity. Method and numerical tests, *Geophys. J. Int.,* **120**(3), 731–744.

Gibbons, S.J. & Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.,* **165**(1), 149–166.

Gibbons, S.J., Kværna, T. & Ringdal, F., 2005. Monitoring of seismic events from a specific source region using a single regional array: a case study, *J. Seismol.,* **9**(3), 277–294.

Gibbons, S.J., Kværna, T. & Ringdal, F., 2010. Considerations in phase estimation and event location using small-aperture regional seismic arrays, *Pure appl. Geophys.,* **167**(4), 381–399.

Gibbons, S.J., Harris, D.B., Dahl-Jensen, T., Kværna, T., Larsen, T.B., Paulsen, B. & Voss, P.H., 2017. Locating seismicity on the Arctic plate boundary using multiple-event techniques and empirical signal processing, *Geophys. J. Int.,* **211**(3), 1613–1627.

Gilbert, F., 1971. Excitation of the normal modes of the earth by earthquake sources, *Geophys. J. Int.,* **22**(2), 223–226.

Grant, M. & Boyd, S., 2008. in Blondel, V., Boyd, S. & Kimura, H., eds, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control,* Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited.

Grant, M. & Boyd, S., 2014. *CVX: Matlab software for disciplined convex programming, version 2.1*. Available at: http://cvxr.com/cvx, (Last accessed 2018 July 12).

Gualtieri, L. & Ekström, G., 2016. Seismic reconstruction of the 2012 Palisades rockfall using the analytical solution to Lamb's problem, *Bull. seism. Soc. Am.,* **107**(1), 63–71.

Gualtieri, L. & Ekström, G., 2018. Broadband seismic analysis and modeling of the 2015 Taan Fjord, Alaska landslide using Instaseis, *Geophys. J. Int.,* **213**(3), 1912–1923,

Hauksson, E. & Shearer, P., 2005. Southern California hypocenter relocation with waveform cross-correlation, Part 1: Results using the double-difference method, *Bull. seism. Soc. Am.,* **95**(3), 896–903.

Hauksson, E., Stock, J., Hutton, K., Yang, W., Vidal-Villegas, J.A. & Kanamori, H., 2011. The 2010 Mw 7.2 El Mayor-Cucapah Earthquake Sequence, Baja California, Mexico and Southernmost California, USA: active Seismotectonics along the Mexican Pacific Margin, *Pure appl. Geophys.,* **168**(8), 1255–1277.

Havskov, J. & Ottemöller, L., 2010. *Array Processing,* Routine Data Processing in Earthquake Seismology Springer, .

Hedlin, M.A.H., de Groot-Hedlin, C.D., Forbes, J.M. & Drob, D.P. 2018. Solar terminator waves in surface pressure observations, *Geophys. Res. Lett.,* **45,** 5213–5219.

Hedlin, M.A., Minster, J.B. & Orcutt, J.A., 1991. Beam-stack imaging using a small aperture array, *Geophys. Res. Lett.,* **18**(9), 1771–1774

Hibert, C., Ekström, G. & Stark, C.P., 2014. Dynamics of the Bingham Canyon Mine landslides from seismic signal analysis, *Geophys. Res. Lett.,* **41**(13), 4535–4541.

Hibert, C., Stark, C.P. & Ekström, G., 2015. Dynamics of the Oso-Steelhead landslide from broadband seismic analysis, *Nat. Hazards Earth Syst. Sci.,* **15**(6), 1265–1273.

Huang, Y. & Beroza, G.C., 2015. Temporal variation in the magnitude-frequency distribution during the Guy-Greenbrier earthquake sequence, *Geophys. Res. Lett.,* **42**(16), 6639–6646.

International Seismological Centre, 2013. *On-line Bulletin,* Int. Seis. Cent., Thatcham, UK. Available at: http://www.isc.ac.uk, (Last accessed 2018 July 12).

Iverson, R. *et al.*, 2015. Landslide mobility and hazards: implications of the 2014 Oso disaster, *Earth planet. Sci. Lett.,* **412,** 197–208.

Ji, C., Tsuboi, S., Komatitsch, D. & Tromp, J., 2005. Rayleigh-wave multipathing along the west coast of North America, *Bull. seism. Soc. Am.,* **95**(6), 2115–2124.

Jordan, T.H. & Sverdrup, K.A., 1981. Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. seism. Soc. Am.,* **71**(4), 1105–1130.

Joughin, I. *et al.*, 2008. Ice–front variation and tidewater behavior on Helheim and Kangerdlugssuaq Glaciers, Greenland, *J. geophys. Res.,* **113**(F1), doi:10.1029/2007JF000837.

Larson, E.W. & Ekström, G., 2002. Determining surface wave arrival angle anomalies, *J. geophys. Res.,* **107**(B6), doi:10.1029/2000JB000048.

Laske, G., 1995. Global observation of off-great-circle propagation of long-period surface waves, *Geophys. J. Int.,* **123**(1), 245–259.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.,* **101**(B7), 16059–16075.

Laske, G. & Masters, G., 1998. Surface-wave polarization data and global anisotropic structure, *Geophys. J. Int.,* **132**(3), 508–520.

Lee, D.-T. & Schachter, B.J., 1980. Two algorithms for constructing a Delaunay triangulation, *Int. J. Comput. Inform. Sci.,* **9**(3), 219–242.

Li, Z., Peng, Z., Hollis, D., Zhu, L. & McClellan, J., 2018. High-resolution seismic event detection using local similarity for Large-N arrays, *Sci. Rep.,* **8**(1), 1646, doi:10.1038/s41598-018-19728-w.

Lin, G., 2013. Three–dimensional seismic velocity structure and precise earthquake relocations in the Salton trough, southern California, *Bull. seism. Soc. Am.,* **103**(5), 2694–2708.

Lin, G. & Shearer, P., 2006. The COMPLOC earthquake location package, *Seismol. Res. Lett.,* **77**(4), 440–444.

Lin, G., Shearer, P.M. & Hauksson, E., 2007. Applying a three–dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. geophys. Res.,* **112**(B12), B12309, doi:10.1029/2007JB004986.

Ma, Z. & Masters, G., 2015. Effect of earthquake locations on Rayleigh wave azimuthal anisotropy models, *Geophys. J. Int.,* **203**(2), 1319–1333.

Ma, Z., Masters, G., Laske, G. & Pasyanos, M., 2014. A comprehensive dispersion model of surface wave phase and group velocity for the globe, *Geophys. J. Int.,* **199**(1), 113–135.

Ma, Z., Masters, G. & Mancinelli, N., 2015. Two-dimensional global Rayleigh wave attenuation model by accounting for finite-frequency focusing and defocusing effect, *Geophys. J. Int.,* **204**(1), 631–649.

Matoza, R.S., Shearer, P.M., Lin, G., Wolfe, C.J. & Okubo, P.G., 2013. Systematic relocation of seismicity on Hawaii Island from 1992 to 2009 using waveform cross correlation and cluster analysis, *J. geophys. Res.,* **118**(5), 2275–2288.

McGuire, J.J., 2008. Seismic cycles and earthquake predictability on East Pacific Rise transform faults, *Bull. seism. Soc. Am.,* **98**(3), 1067–1084.

Meng, X. & Peng, Z., 2014. Seismicity rate changes in the Salton Sea Geothermal Field and the San Jacinto Fault Zone after the 2010 Mw 7.2 El Mayor-Cucapah earthquake, *Geophys. J. Int.,* **197**(3), 1750–1762.

Murray, T. *et al.*, 2015a. Reverse glacier motion during iceberg calving and the cause of glacial earthquakes, *Science,* **349,** 305–308.

Murray, T. *et al.*, 2015b. Dynamics of glacier calving at the ungrounded margin of Helheim Glacier, southeast Greenland, *J. geophys. Res.,* **120**(6), 964–982.

Nettles, M. & Ekström, G., 2010. Glacial earthquakes in Greenland and Antarctica, *Annu. Rev. Earth Planet. Sci.,* **38**(1), 467–491.

Nettles, M. *et al.*, 2008. Step-wise changes in glacier flow speed coincide with calving and glacial earthquakes at Helheim Glacier, Greenland, *Geophys. Res. Lett.,* **35**(24), doi:10.1029/2008GL036127.

Olsen, K.G. & Nettles, M., 2017. Patterns in glacial-earthquake activity around Greenland, 2011–13, *J. Glaciol.,* **63**(242), 1077–1089.

Pedersen, H., Boué, P., Poli, P. & Colombi, A., 2015. Arrival angle anomalies of Rayleigh waves observed at a broadband array: a systematic study based on earthquake data, full waveform simulations and noise correlations, *Geophys. J. Int.,* **203**(3), 1626–1641.

Poli, P., 2017. Creep and slip: seismic precursors to the Nuugaatsiaq landslide (Greenland), *Geophys. Res. Lett.,* **44**(17), 8832–8836.

Ringdal, F. & Husebye, E.S., 1982. Application of arrays in the detection, location, and identification of seismic events, *Bull. seism. Soc. Am.,* **72**(6B), S201–S224.

Ruan, Y. & Zhou, Y., 2012. The effects of 3-D anelasticity (Q) structure on surface wave amplitudes, *Geophys. J. Int.,* **189**(2), 967–983.

Ruud, B.O., Husebye, E.S., Ingate, S.F. & Christoffersson, A., 1988. Event location at any distance using seismic data from a single, three-component station, *Bull. seism. Soc. Am.,* **78**(1), 308–325.

Shearer, P., Hauksson, E. & Lin, G., 2005. Southern California hypocenter relocation with waveform cross–correlation, Part 2: results using source-specific station terms and cluster analysis, *Bull. seism. Soc. Am.,* **95**(3), 904–915.

Shearer, P.M., 1994. Global seismic event detection using a matched filter on long–period seismograms, *J. geophys. Res.,* **99**(B7), 13713–13725.

Shelly, D.R. & Hardebeck, J.L., 2010. Precise tremor source locations and amplitude variations along the lower–crustal central San Andreas Fault, *Geophys. Res. Lett.,* **37**(14), doi:10.1029/2010GL043672.

Tanimoto, T. & Prindle, K., 2007. Surface wave analysis with beamforming, *Earth Planets Space,* **59**(5), 453–458.

Thompson, C. & Shure, L., 2016. *Image Processing Toolbox: For Use with MATLAB; [User's Guide],* MathWorks.

Thurber, C.H., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. geophys. Res.,* **88**(B10), 8226–8236.

Thurber, C.H., 1985. Nonlinear earthquake location: theory and examples, *Bull. seism. Soc. Am.,* **75**(3), 779–790.

Trugman, D.T. & Shearer, P.M., 2017. Growclust: a hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences, *Seismol. Res. Letters,* **88**(2A), 379–391.

Tsai, V.C. & Ekström, G., 2007. Analysis of glacial earthquakes, *J. geophys. Res.,* **112**(F3), doi:10.1029/2006JF000596.

Tsai, V.C., Rice, J.R. & Fahnestock, M., 2008. Possible mechanisms for glacial earthquakes, *J. geophys. Res.,* **113**(F3), doi:10.1029/2007JF000944.

Vandenberghe, L. & Boyd, S., 1996. Semidefinite programming, *SIAM Rev.,* **38**(1), 49–95.

Vasco, D., Johnson, L.R. & Majer, E., 1993. Ensemble inference in geophysical inverse problems, *Geophys. J. Int.,* **115**(3), 711–728.

Veitch, S.A. & Nettles, M., 2012. Spatial and temporal variations in Greenland glacial-earthquake activity, 1993–2010, *J. geophys. Res.,* **117**(F4), doi:10.1029/2012JF002412.

Waldhauser, F. & Ellsworth, W.L., 2000. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California, *Bull. seism. Soc. Am.,* **90**(6), 1353–1368.

Waldhauser, F. & Schaff, D.P., 2008. Large-scale relocation of two decades of northern California seismicity using cross-correlation and double-difference methods, *J. geophys. Res.,* **113**(B8), B08311, doi:10.1029/2007JB005479.

Wei, S. *et al.*, 2011. Superficial simplicity of the 2010 El Mayor-Cucapah earthquake of Baja California in Mexico, *Nat. Geosci.,* **4**(9), 615–618.

Woodhouse, J. & Wong, Y., 1986. Amplitude, phase and path anomalies of mantle waves, *Geophys. J. Int.,* **87**(3), 753–773.

Yoon, C.E., O'Reilly, O., Bergen, K.J. & Beroza, G.C., 2015. Earthquake detection through computationally efficient similarity search, *Sci. Adv.,* **1**(11), doi:10.1126/sciadv.1501057.

Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.,* **5**(7), 493–498.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Table S1. Detected seismic sources.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.